Iranian Society of Acoustics and Vibration
The 13th ISAV2023 International Conference on Acoustics and Vibration
20, 21 Dec 2023 Tehran - Iran

Leveraging Koopman Operators and Deep Neural Networks for Parameter Estimation and Future Prediction of Duffing Oscillators

Yassin Riyazi [a],NavidReza Ghanbari [a], Arash Bahrami [a*]

[a] M.Sc. Student, School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran.

[b] Assistant Professor, School of Mechanical Engineering, College of Engineering, University of Tehran, Tehran, Iran.

** Corresponding author e-mail: arash.bahrami@ut.ac.ir (A. Bahrami)*

Abstract

The study of nonlinear dynamical systems has been a cornerstone in various scientific and engineering domains due to their widespread applicability in modeling real-world phenomena. Traditional methods for analyzing and predicting the behavior of such systems often involve complex mathematical techniques and numerical simulations. This paper introduces a novel approach that combines the power of Koopman Operators and deep neural networks to generate a linear representation of the Duffing oscillator, enabling effective parameter estimation and accurate prediction of its future behavior. Furthermore, a modified loss function is proposed to enhance the training process of the deep neural network. The synergy of Koopman Operators and deep neural networks not only simplifies the analysis of nonlinear systems but also offers a promising avenue for advancing predictive modeling in various fields.

Keywords: Koopman Operator; Parameter Estimation; Koopman Operator; Duffing oscillator; deep neural networks; parameter estimation; nonlinear dynamical systems; loss function; predictive modeling.

1. Introduction

Nonlinear dynamical systems, celebrated for their intricate and often chaotic behavior, permeate the realms of natural phenomena and technological applications. They transcend the simplicity of linear systems, giving rise to phenomena such as bifurcations, limit cycles, and chaotic attractors. These systems have long captivated the interest of scientists and engineers, presenting substantial challenges in understanding, characterizing, and predicting their trajectories. Across diverse domains, from physics and biology to economics and engineering, nonlinear systems underscore the fundamental complexity of our world.

At the heart of this intricate landscape lies the Duffing Oscillator [1], an iconic archetype of nonlinear dynamical systems. Its versatility enables it to emulate a wide spectrum of behaviors, making it a pertinent model for various physical phenomena. From capturing the subtle interplay of mechanical vibrations in structures subjected to external forces to mirroring the rhythmic patterns of biological oscillations, the Duffing oscillator encapsulates the essence of nonlinear dynamics.

Traditionally, dissecting and forecasting the behavior of Duffing oscillators has relied on a combination of analytical techniques and numerical simulations. While these methods provide valuable insights, they often encounter limitations in handling nonlinear intricacies with precision. Analytical solutions may prove elusive or computationally expensive, especially for higher-dimensional or strongly nonlinear systems. Conversely, numerical simulations, though powerful, demand extensive computational resources and face challenges in long-term predictions due to inherent numerical errors and uncertainties.

To address these challenges, we introduce an innovative approach that harnesses the synergy between Koopman Operators [2] and Deep Neural Networks. [3]–[5] This groundbreaking fusion aims to transform the Duffing oscillator into a linearized representation, offering promising solutions to the complexities posed by traditional methods. By leveraging the power of deep learning and the Koopman operator's ability to provide a linear representation of nonlinear systems [6], our approach enables a more accurate Koopman linearized representation of system behavior.

Our approach not only simplifies the analysis of nonlinear systems but also extends its applicability to a diverse range of domains. It heralds a new era in predictive modeling by opening avenues for streamlined parameter estimation and accurate future predictions, addressing the challenges posed by the inherent complexity of nonlinear systems.

Brunton's group has been incorporating autoencoder structure [7] with SINDY method [8] and successfully identify the system [9][10].

In the following sections, we delve into the foundational principles of Koopman Operator Theory and the adaptability of deep neural networks. We showcase how their fusion forms a compelling framework for analyzing and predicting the behavior of Duffing oscillators. We outline the process of transforming Duffing oscillator dynamics into a linear representation and introduce a modified loss function designed to enhance the generality of the Koopman linear representation of the dynamical system within this context. Through experimental validation and comparisons with traditional methods, we demonstrate the efficacy of our approach in providing accurate predictions for the future behavior of Duffing oscillators. Ultimately, this work enriches our understanding of nonlinear dynamics and offers a powerful toolset with transformative potential across scientific, engineering, and practical applications.

2. Koopman Operator and Its Application:

The exploration of dynamical systems has long been a cornerstone in understanding complex behaviors in various scientific and engineering disciplines. Traditionally, the analysis of these systems has been deeply entwined with the concept of state space, where the evolution of a system is represented by trajectories in the space of its state variables. However, the inherent nonlinearity of many real-world systems often makes their analysis and prediction challenging using conventional techniques.

In recent years, the Koopman Operator has emerged as a powerful mathematical tool that provides a fresh vantage point for studying dynamical systems. Rooted in functional analysis, the

Koopman Operator introduces a paradigm shift by transitioning the focus from the state space to the space of observable functions. Doing so allows us to view the system's evolution in a linear framework, even when dealing with inherently nonlinear systems. This perspective offers a new lens through which we can gain insights into the dynamics of complex systems.

2.1 Dynamical System Representation

Consider a dynamical system described by a set of state variables $\mathbf{x}(t)$, which evolve over time t. Mathematically, we can represent this as:

$$\frac{d\mathbf{x}}{dt} = f(\mathbf{x}, t) \tag{1}$$

Where:

- n is the state vector dimension.
- $\mathbf{x} \in \mathbb{R}^n$ is the state vector representing the system's state variables.
- $t \in \mathbb{R}^+$ is time.
- $f(\mathbf{x}) : \mathbb{R}^n \to \mathbb{R}^n$ is a vector-valued function describing how the state variables change over time.

2.2 Koopman Operator Transformation

The Koopman operator, denoted as $\mathcal{K}$, is an infinite-dimensional linear operator that acts on observables or functions of the state variables. Let $g(x)$ be such an observable. The Koopman operator maps this observable from the state space to a higher-dimensional space:

$$\mathcal{K}g(\mathbf{x}) = g(f(\mathbf{x})) \tag{2}$$

Where:

- m supposed to be infini-dimensional but in numerical approximation a value will be assigned.
- $\mathcal{K}(.) : \mathbb{C}^m \to \mathbb{C}^m$ is the Koopman operator.
- $g(\mathbf{x}) : \mathbb{R}^n \to \mathbb{C}^m$ is an observable or function defined on the state space.
- $g(f(\mathbf{x}))$ represents the observable after the system evolves according to $f(\mathbf{x})$.

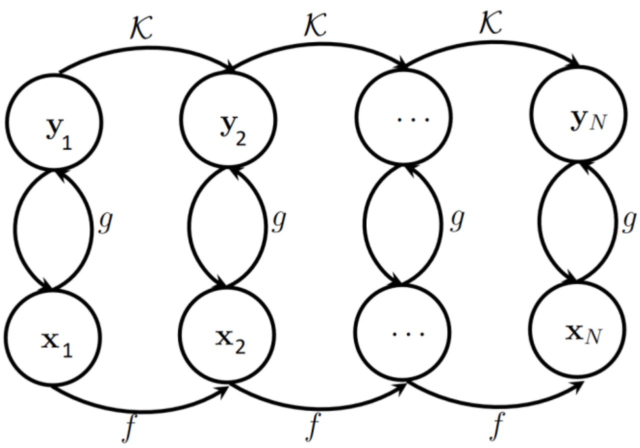

Figure 1. Koopman Operator Evolution and a discrete dynamical system

2.3 Koopman Operator in Discrete Time

In many cases, you may work with discrete-time dynamical systems. In this case, the Koopman operator is applied at discrete time steps. The equation becomes:

$$\mathcal{K}g\left(\mathbf{x}_k\right)=g\left(\mathbf{x}_{k+1}\right) \tag{3}$$

Where:

- $\mathbf{x}_k$ represents the state of the system at time k.
- $\mathbf{x}_{k+1}$ represents the state of the system at the next time step $k+1$.

3. Coupling Koopman Operators with Deep Neural Networks

Deep neural networks have showcased remarkable abilities in approximating intricate functions and mastering complex patterns from data. One of the key challenges encountered in the realm of Koopman Operators is the identification of suitable observable functions. In methods such as DMD [11], the observable function is typically the identity function, and in extended DMD (EDMD), observable functions take the form of polynomials or trigonometric functions. While these approaches are straightforward and accurate, they exhibit resilience to noise and initial conditions.

In this study, we amalgamate a deep neural network with the Koopman operator, thereby generating a linearized representation of the Duffing oscillator. This neural network effectively learns the intricate relationship between system parameters and observed behaviors, facilitating efficient parameter estimation. Moreover, the neural network undergoes training to predict the future trajectory of the Duffing oscillator, thereby equipping us with a valuable tool for forecasting system behavior.

neural networks aren't necessarily always better than feature crosses, but neural networks do offer a flexible alternative that works well in many cases.

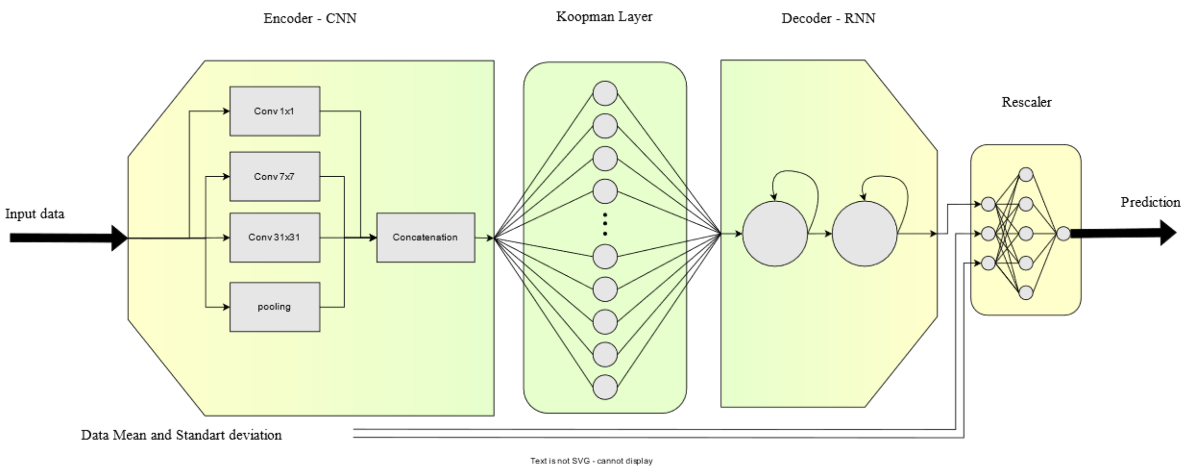

Figure 2. Neural Network diagram

3.1 Structure

The architecture of the model as shown is Figure 2. Neural Network diagram is structured around an Encoder-Decoder paradigm, which effectively captures the essence of complex dynamics.

Specifically, the Encoder component is meticulously designed, featuring a sequence of Inception Blocks [12] in a convolutional neural network (CNN) [13]. These Inception Blocks serve as robust feature extractors, enabling the model to discern intricate patterns and relevant features from the input data.

Following the Inception-based Encoder, a pivotal transformation takes place through a linear layer. This linear layer assumes a distinct role within the architecture, embodying the essence of the Koopman Operator evolution function $\mathcal{K}$. It's important to note that this linear layer operates without an activation function and bias, preserving the linear nature of the Koopman operator's transformation.

Transitioning from the Koopman Operator layer, the architecture takes an intriguing turn with the integration of a two-layer Long Short-Term Memory (LSTM) [14] network. This LSTM component acts as the Decoder, expertly leveraging its sequential memory to unravel the transformed linearized representation. This sequence-to-sequence modeling approach facilitates the reconstruction of the system's temporal evolution, a crucial aspect in capturing its intricate behaviors.

Table 1. CNN's parameters. Out hyper parameter is 20

	Sequential Blocks	In_Channels	Out_Channels	kernel size	padding
branch1x1	Conv 1×1	1	out	1	0
	ReLU	-	-	-	-
branch7x7	Conv 1×1	1	out	1	0
	ReLU	-	-	-	-
	Conv 7×7	out	out	7	3
	ReLU	-	-	-	-
branch31x31	Conv 1×1	1	out	1	0
	ReLU	-	-	-	-
	Conv 31×31	out	out		15
	ReLU	-	-	-	-
branch_pool	MaxPool1d	1	1	3	1
	-	-	-	-	-
	Conv 1×1	1	out	1	0
	ReLU	-	-	-	-

Add table of RNN's

3.2 Training

The network is trained end-to-end, without the need for custom loss functions or specialized training algorithms. However, it's important to note that the evolution function of the Koopman operator doesn't remain confined solely to the Koopman part; instead, it spreads throughout the network. In a sense, the network operates as a black box, handling this evolution internally.

To address this issue and restrict the Koopman Operator's influence exclusively to the Koopman linear layer, a two-stage training algorithm has been proposed. In this algorithm, after each optimization step:

1. The weights of all layers except the Koopman Linear Layer are frozen.
2. The output of the Koopman Linear Layer is calculated for time steps n_0 to n_{KPH} (KPH is the a Hyper parameters and due to cost of calculating matrix power 20 was selected).
3. The weights of the Koopman Linear Layer are updated based on the linearity property. This update aims to minimize the prediction error of the nth output .

$$\sum_{n}^{KPH}\mathcal{L}\left(g\left(x_{n_0}\right)\times\left(W_{Koopman}^{n}\right)^{T},g\left(x_n\right)\times\left(W_{Koopman}\right)^{T}\right) \tag{4}$$

By implementing this two-stage training process, we ensure that the Koopman Operator's influence is confined and utilized specifically within the Koopman linear layer, enhancing the network's predictive accuracy and control.

Table 2. Optimizer and loss parameters

		Stage 1	Stage 2
Optimizer	Type	SGD	SGD
	Learning Rate	5.00E-02	5.00E-04
	momentum	0.9	-
	weight_decay	1.00E-04	-
loss	Type	MSE	MSE
	alpha	0	0

3.3 Loss Function

Due to intrinsic characteristic of the numerical job and unfeasibility of the infinite-dimensional Hilbert space to ensuring the Koopman Operator is full rank , a modified loss function is proposed. Traditional loss functions may not assure the Koopman Generator Matrix is Fully ranked the intricacies of nonlinear dynamics, leading to suboptimal training outcomes. By tailoring the loss function to align with the unique characteristics of the Duffing oscillator, the network's performance is significantly improved, resulting in more accurate parameter estimates and predictions.

3.4 Data Normalization

Given the nature of regression, it is advisable to normalize the data before training the neural network. This normalization is crucial because even small variations in the input data can lead to significant changes in the output, potentially reducing the model's robustness against changes in input conditions. In our approach, we employ data normalization prior to feeding it into the convolutional neural network (CNN) architecture. During this process, we pass the statistical properties of the data, such as its mean and variance, through the network.

In Figure 2. Neural Network diagram, after the recurrent neural network (RNN) block, the data is remapped to its original statistical properties before being passed to the Rescaler Block. The purpose of the Rescaler Block is to further reduce the variation in the output, ultimately leading to a more stable and controlled model response.

3.5 Data acquisition

The Duffing oscillator is a dynamical system described by the following second-order differential equation:

$$\frac{d^2x}{dt^2}+\delta\frac{dx}{dt}+\alpha x+\beta x^3=\gamma\cos\left(\omega t\right) \tag{4}$$

Here:

- x represents the displacement of the oscillator from its equilibrium position.
- δ denotes the damping coefficient.
- α is the linear stiffness coefficient.
- β characterizes the nonlinearity in the system.
- γ is the amplitude of the external driving force.

- ω is the angular frequency of the driving force.
-

Duffing oscillator solution has been generated using Runge-Kuta method [reference] and initial condition for Solving the equation is $x_0 = 1.5$ and $v_0 = -1.5$.

A normal distributes with nose in range of [-0.5,0.5] added to data to simulate real world data.

4. Results and Discussion

Experimental results demonstrate the effectiveness of the proposed approach. The combination of Koopman operator-based linearization and deep neural networks yields impressive results in terms of parameter estimation accuracy and future prediction. The modified loss function further enhances the network's training process, underscoring its importance in capturing the nuances of nonlinear dynamics.

Neural network is robust against noise with range (-0.5,0.5) to (-0.07,0.07)

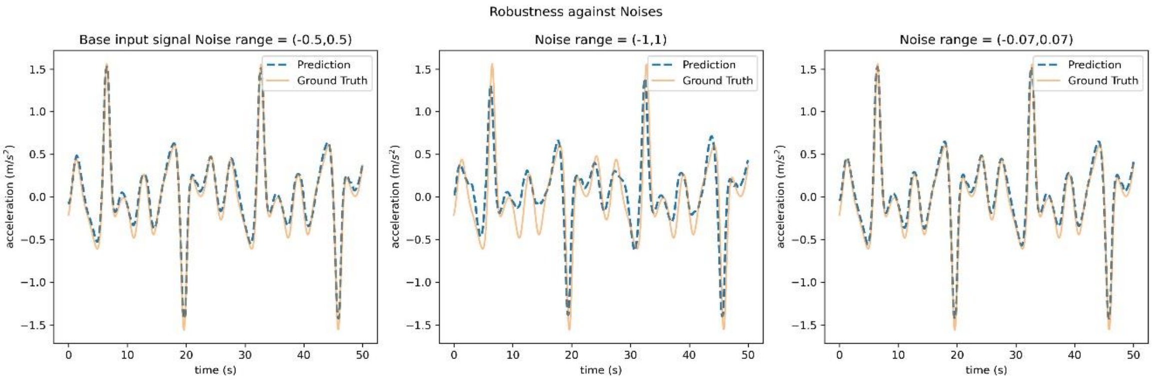

Figure 3. robust against noise

When considering various initial conditions while keeping the parameters (gammas) constant, the neural networks perform adequately. However, it is important to note that ensuring the networks do not achieve a loss lower than 0.01 is advisable, as excessively low losses may increase the risk of overfitting.

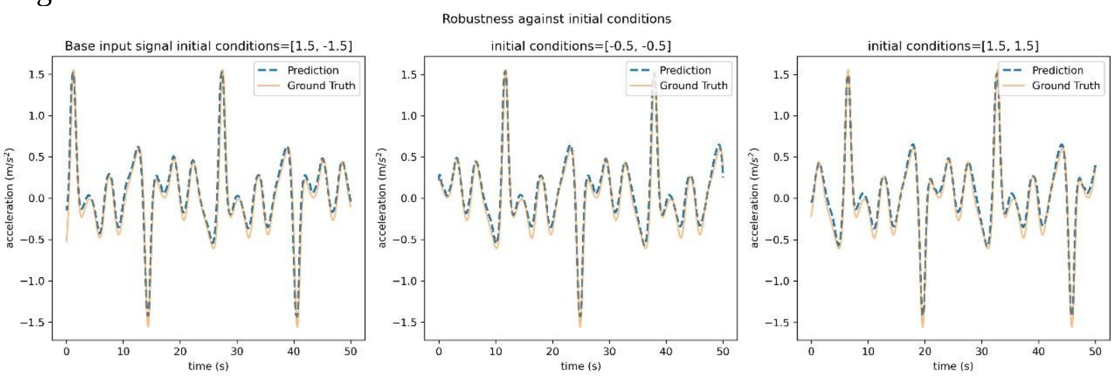

Figure 4. robust against initial condition

Different values of gammas do not yield satisfactory results, as datasets corresponding to distinct gammas exhibit entirely different structures. Introducing diverse gamma values during training may also lead to suboptimal outcomes. An alternative approach worth exploring is the expansion of the network's capacity, which could potentially enhance its performance under such circumstances.

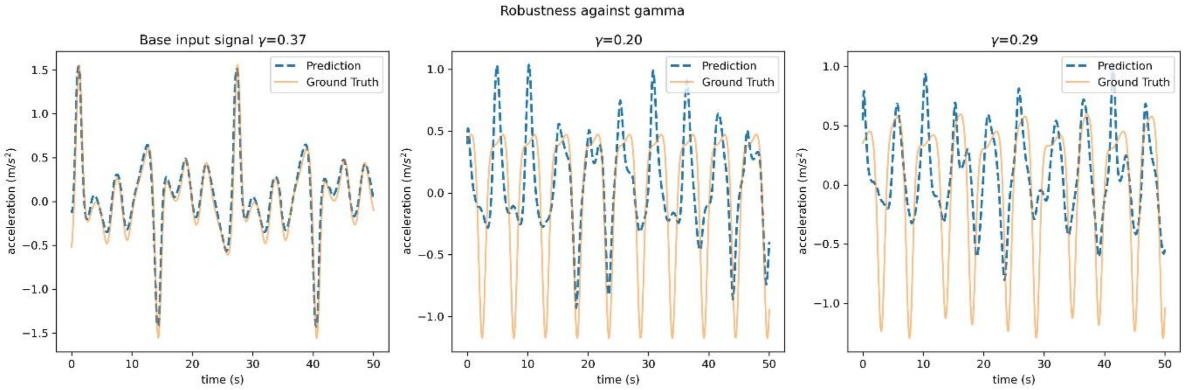

Figure 5. robust against γ

Various sampling rates yielded suboptimal results as well.

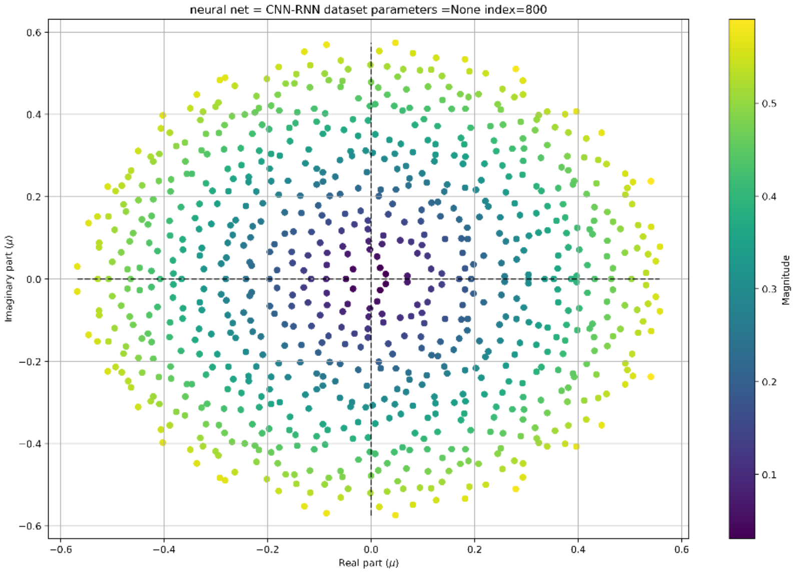

Figure 6. Koopman Layer Eigen values

5. Conclusion

In a broader context, the combination of the Koopman Operator and Neural Networks shows significant potential. The network has effectively captured the underlying data structure, and with further refinement, it has the capacity to generalize effectively to more complex problem domains.

6. Future studies

In the future research endeavors, several promising avenues can be explored:

1. Transitioning from Offline to Online Processing: An interesting prospect is the shift from offline to online data processing. Implementing real-time data analysis and prediction systems can enhance the applicability and timeliness of the models developed.
2. Integration of Gradient Clipping: The inclusion of gradient clipping techniques in training neural networks should be considered. This can help mitigate issues related to exploding gradients and improve model stability during training.
3. Incorporation of Mixture Density Networks for Confidence Estimation: Particularly in scenarios involving chaotic dynamics, the addition of Mixture Density Networks (MDNs) at

the model's output can provide valuable confidence estimates. This can enhance the reliability of predictions, especially when dealing with inherently uncertain or complex data.

4. Exploration of Small Transformer Models: The application of smaller-scale Transformer models warrants investigation. These models can serve a dual purpose: identifying any potentially missed data patterns within signals and subsequently utilizing these identified patterns for prediction tasks or noise reduction, thus improving overall data quality.

These avenues represent promising directions for advancing the research in this domain, with the potential to yield enhanced model performance and broader applications.

- Make sense of using Convolutions in Koopman Operator and prove that is a nonlinear function like Brunton's research
- Compare it with ARMA,ARMAX and other estimation methods
- Bring the Config of the Computer and the Flops required to train
- تبدیل معکوس تنها برای جواب های رو منیفلد کار میکند و وقتی نویز داریم دیگر تبدیل معکوس یکتا وجود ندارد و مهم نیست تابع واران چه باشد. ارجاع بده به شبکه ای که علی دایت ارایه داد

REFERENCES

[1] Hamel, "Georg Duffing, Ingenieur: Erzwungene Schwingungen bei veränderlicher Eigenfrequenz und ihre technische Bedeutung. Sammlung Vieweg. Heft 41/42, Braunschweig 1918. VI+134 S," *ZAMM - J. Appl. Math. Mech. / Zeitschrift für Angew. Math. und Mech.*, vol. 1, no. 1, 1921, doi: 10.1002/zamm.19210010109.

[2] B. O. Koopman, "Hamiltonian Systems and Transformation in Hilbert Space," *Proc. Natl. Acad. Sci.*, vol. 17, no. 5, 1931, doi: 10.1073/pnas.17.5.315.

[3] A. Krizhevsky, I. Sutskever, and G. E. Hinton, "2012 AlexNet," *Adv. Neural Inf. Process. Syst.*, 2012.

[4] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," in *3rd International Conference on Learning Representations, ICLR 2015 - Conference Track Proceedings*, 2015.

[5] M. Shafiq and Z. Gu, "Deep Residual Learning for Image Recognition: A Survey," *Applied Sciences (Switzerland)*, vol. 12, no. 18. 2022. doi: 10.3390/app12188972.

[6] S. L. Brunton, M. Budišić, E. Kaiser, and J. N. Kutz, "Modern Koopman Theory for Dynamical Systems," *SIAM Rev.*, vol. 64, no. 2, 2022, doi: 10.1137/21M1401243.

[7] G. E. Hinton and R. R. Salakhutdinov, "Reducing the dimensionality of data with neural networks," *Science (80-.).*, vol. 313, no. 5786, 2006, doi: 10.1126/science.1127647.

[8] S. L. Brunton, J. L. Proctor, and J. N. Kutz, "Discovering governing equations from data by sparse identification of nonlinear dynamical systems," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 113, no. 15, 2016, doi: 10.1073/pnas.1517384113.

[9] K. Champion, B. Lusch, J. Nathan Kutz, and S. L. Brunton, "Data-driven discovery of coordinates and governing equations," *Proc. Natl. Acad. Sci. U. S. A.*, vol. 116, no. 45, 2019, doi: 10.1073/pnas.1906995116.

[10] B. Lusch, J. N. Kutz, and S. L. Brunton, "Deep learning for universal linear embeddings of nonlinear dynamics," *Nat. Commun.*, vol. 9, no. 1, 2018, doi: 10.1038/s41467-018-07210-0.

[11] P. J. Schmid, "Dynamic mode decomposition of numerical and experimental data," *J. Fluid Mech.*, vol. 656, 2010, doi: 10.1017/S0022112010001217.

[12] C. Szegedy *et al.*, "Going deeper with convolutions," in *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 2015. doi: 10.1109/CVPR.2015.7298594.

[13] Y. LeCun, L. Bottou, Y. Bengio, and P. Haffner, "Gradient-based learning applied to document recognition," *Proc. IEEE*, vol. 86, no. 11, 1998, doi: 10.1109/5.726791.

[14] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory," *Neural Comput.*, vol. 9, no. 8, 1997, doi: 10.1162/neco.1997.9.8.1735.